Chapter 6

Patterns of Quality of Life in Facilities

Once we developed scales to use to characterize an individual resident's quality of life (QOL), we could explore additional questions, particularly: what characteristics are associated with residents reporting a better or worse QOL on various domains; and to what extent are nursing homes distinguished from each other by the QOL that their residents tend to report. For regulatory purposes, it may be important to determine that facilities do indeed differ in the average reported QOL of their residents. For a nursing home's own continuous quality improvement efforts, information about correlates of QOL at the individual level could help target the efforts.

Background

Central to all facility-level uses of the QOL data is an assumption that nursing homes should be held accountable for QOL outcomes of interest. Nursing homes differ from many other health care programs in that, for many residents, they serve as dwelling places for long periods of time. The nursing home, therefore, has the potential to influence residents' lives for better or worse on outcomes that are far broader than health status. Arguably, the nursing home's effects on its residents' QOL are at least as important as the more traditional measures of quality of care that are routinely collected.

Some controversy exists, as yet uninformed by much data, about the extent to which a nursing home can actually influence the social and psychological domains of QOL. For example, outcomes such as meaningful activity, relationships, and the like may be heavily determined by social factors outside facility control, such as family structure and availability, quality of family relationships, and resident's interests, education, and even personality. It is also

reasonable to hypothesize that QOL levels in a nursing home will be related to various health and disability characteristics such as: health status, prognosis, functional abilities, sensory abilities, pain, persistent serious depressive illness, and cognitive abilities. These factors can affect how a resident appraises his or her QOL. Nursing homes, however, can likely influence some of these potential mediators of QOL directly, especially pain, functional status, and depression (particularly depressive affect that is not part of a longstanding psychiatric diagnosis). Although it is difficult for nursing homes to overcome the effects of extreme sensory impairment, high disease burden, irreversible cognitive impairment, and poor prognoses, we suggest that nursing homes can take effective steps to improve or maintain QOL for those who may be at highest risk of poor QOL, including those with dementia, those who are facing imminent death, and those who have limited external social support systems.

Indicators of quality of a health-care organization may be expressed at several levels. One way to generate indicators is to identify facility-level structural and process factors that relate to outcomes of interest, such as nursing staff levels or infection control. Another approach is to aggregate individual level data on outcomes of interest (e.g. proportion of bedsores, proportion of urinary tract infections) to describe the facility. When such outcomes are aggregated, adjustments are necessary for aspects of the case-mix that are out of the control of the facility. Using the second approach, nursing home data generated through the Minimum Data Set (MDS) have been under study for more than a decade and various health-oriented outcome indicators have been developed using different case-mix adjustment strategies. For quality assurance purposes, and for presenting meaningful comparative data to the public, it is necessary to avoid describing nursing homes as better in quality if a substantial amount of the differences can be attributed to the characteristics of the residents who are admitted. On the other hand, one must

not over-adjust and, therefore, fail to hold nursing homes accountable for characteristics of residents that they may be able to change (e.g. bed-bound, wheelchair-bound residents, or residents who are depressed).

A central task in creating a facility-level measure of QOL is to adjust appropriately for differences in case mix across facilities, although this need is not always appreciated (Davis, 1991). The choice of case-mix adjusters is important; they should reflect elements that might influence QOL, but they should not include items that are under the control of the nursing home, lest important differences in quality are adjusted away. Although a growing body of information is available on case-mix adjustment for quality of care in nursing homes (Braun, 1991) (Mukamel, 1997) (Phillips et al., 1996; Porell & Caro, 1998), little work has been done on QOL.

Methods

Data

Data for these analyses are derived from the resident interviews in the two waves of data collection, information abstracted from each resident's chart (e.g., birth date, date of admission), and data derived from the MDS. In addition to QOL data, the resident interview also collected basic demographic data (e.g., gender, race, marital status). We had acquired the MDS files for all the residents in all the 101 participating facilities for 18 months prior to our data collection for both Wave1 and Wave 2. This permitted us to use individual information on each resident in the sample for adjustment purposes; for example, we used data from the cognition and ADL sections for such adjustments. It also permitted us to create facility averages of case-mix characteristics of interest. For adjustment purposes, we used the MDS evaluation closest to the time of the resident interview to create the cognitive and ADL adjustment. The latter was based on a count of the number of ADLs in which the resident was rated as requiring assistance.

Finally, facility descriptors such as size, urban or rural location were drawn from our own date base.

Analysis

The comparison of QOL among facilities was implemented on two levels: raw scores and scores adjusted for differences in patient characteristics. First, QOL scores were calculated using all available resident-level responses and aggregated by facilities to obtain non-adjusted, facility-level average QOL scores. Resident-level data were also used to develop a case-mix adjusted model that regressed individual QOL scores on selected resident characteristics obtained from MDS data. These included ADL, cognition, age, gender, and length of stay. Length of stay was dichotomized into less than three months and three months or more.

To enable the comparison of facilities while adjusting for patient characteristics, differences were calculated between the observed individual QOL scores and the expected QOL scores calculated based on the case-mix adjusted model described above. The standardized resident-level residuals were averaged by facility to obtain facility-level adjusted scores. The distribution of the facility-level adjusted scores for all facilities was transformed using z-scores. As a result, the differences among facilities were expressed in units of the standard deviation of average scores for all facilities. By construction (assuming close to a normal distribution), we could expect that approximately half of all facilities would have positive and another half would have negative standardized scores in the range from minus three to plus three.

Two-way analysis of variance was used to compare the differences in distributions among the reports for a given domain within each home and across the homes. Chi square tests were used to compare the proportion of outliers by characteristics of the facilities. Statistical tests were implemented using SPSS 10.1.

Mixed-effect hierarchical linear models with main effects were used to fit the data. Ten components of the QOL instrument were used as dependent variables in Wave 1 and 11 in Wave 2. Independent variables included various combinations of random factors (Facility, Interviewer), and covariates. The design of the models reflected nesting of the data. The allocation of Interviewers by Facilities was not planned in advance and post hoc cross-tabulation revealed very unbalanced design with many empty cells. This imbalance dictated the use of a Type 4 sum of squares that can accommodate the design with empty cells. Covariates included in the models served as risk-adjustors and included length of stay in the nursing facility and MDS-based cognition score (six-level ordinal variable), the ADL score, age and education (five-level ordinal variable), and binary variables representing race, marital status, gender, and the presence of children. These variables were constructed based on the measured values to obtain reasonably unskewed distributions. Calculations were implemented using the General Linear Model procedure in SPSS 11.0.

Results

Wave 1 Findings

Results are presented separately for Waves 1 and 2. The characteristics of the Wave 1 facility sample are shown in Table 6.1, which also contrasts the sample facilities to the national profile on some parameters. The mean size of facilities in the resultant sample was 128 beds; the median was 109.

Table 6.1 Wave 1 Facility Characteristics (N=40)

Characteristics	National Rates [a]	Sample Facility Rates	Source of Data
Mean No. of Beds (Range)	108	128 (49-289)	Initial contact survey
% Urban		50%	Initial contact survey
Mean % single rooms (Range)		24.42% (0,100%)	Initial contact survey
Mean staff/resident ratio (Range)	0.60 – 1.00	0.62 (0.13 – 0.92)[b]	OSCAR
% Proprietary	65.2% (6.7 in AL – 82 in OK)	37.5%	Initial contact survey
Mean % Medicaid residents (Range)	67.7% (49.4 in IA – 84.7 in AL)	67% (0 – 97%)	Initial contact survey
Mean % residents with better cognition—i.e., score of 0-2[c] (Range)		42.6% (11 – 76%)	MDS

Notes:

a. National rates come from 1999 – 2000 (Source: AHCA Facts and Trends: Nursing Facility Sourcebook 2001).
b. Staffing number comprises FTEs of CNAs, LPN/LVNs, and RNs (whether contracted, full time, or part time).
c. Based on the same 6-point MDS-derived cognitive function scale described in Chapter 2 (range from 0-5).

By design, half the sample was rural. There was a considerable range in staffing. The proportion of homes that were proprietary was much lower than the national average, partly because Minnesota and New York have high numbers of nonprofits, but also because most of the refusals occurred in for-profit facilities, whereas all the facilities over-sampled for private rooms were nonprofit. There was considerable variation in the proportion of cognitively impaired residents. This distribution was generally similar across four of the states (averaging about 55%), but California homes had a much higher proportion (67%) than the rest of the sample.

The mean facility scores for each domain are shown in Table 6.2. We created an average score across domains by dividing the additive scale by the number of questions for the particular domain. Each QOL domain could be scored between 4 and 1, with higher scores reflecting

higher QOL. Facility-level scores varied from 2.70 (meaningful activity) to 3.67 (dignity). The extent of between-facility variation in the scores is reflected in the standard deviations, which varies over 100% (from 0.101 for dignity to 0.212 for spiritual well-being).

Table 6.2 QOL Domain Alpha Values and Facility Scores

Domain	Domain Scores		Facility Scores	
	Alpha	No. Items	Mean	Standard Deviation
Comfort	.62	6	3.0210	.1525
Functional Competence	.65	5	3.2463	.1577
Privacy	.52	5	3.3326	.1682
Dignity	.64	5	3.6667	.1011
Meaningful Activity	.77	5	2.6975	.1766
Food Enjoyment	.71	3	3.2176	.1923
Relationships	.70	5	3.0639	.1902
Security	.76	5	3.3939	.1363
Spiritual well-being	.59	4	3.1503	.2120
Autonomy	.64	4	3.2981	.1497

Table 6.3 is designed to show the overall patterns of performance for each facility, by displaying the results arrayed by facility according to whether the facility mean score was above or below the sample average on each domain. To make the distinctions more visible, the facilities are arranged in order of the size of the deviation. The facilities are ranked from best to worse. Each + or - sign represents a standard deviation; the more + or - signs, the greater the facility average deviates from the mean for that domain. The results show that facilities do indeed differ in their average QOL on these domains. In 23 facilities two or more domains were consistently positive or negative by at least one standard deviation. Ten facilities showed a consistently positive pattern and 13 a consistently negative pattern.

Table 6.3. Facility-Level Deviations in Resident-Report QOL (Risk-Adjusted)

NF	Comfort	Functional Competence	Privacy	Dignity	Meaningful Activity	Enjoyment	Relationships	Security	Spiritual well-being	Autonomy
11	+		+		+	+	++		+	
13	+	+				+	+	+		+
28	+		+		+	+		+		+
5		+	+	+					+	
34	+		+	+	+					
6	+			+			+	+		
2		+			+	+				
31		+						+		+
26						+	+			+
1		+			+					
18			+							+
35							+			
32						+				
8								+		--
10			+	-						+
38		+			-	+			+	+
3			+			-	-		-	
4	-			+						
15						+			+	-
36	+	++	-		--	--	-			+
40	+			-	-		-		+	
19	-			+						
20	-	-			+					
30	-						+		+	+
33										
25		-								
37		-								
29		-		-						
7							-		-	
9			-			-			-	
27		-					-			-
23					--	-		---		
24	-		--	---						
16		-				-		-	-	
14					-		-	-	-	
21	-					-	-		-	-
22	-	-		-		-		-		
17	-		-	-				-	-	
39	-		-		-		-		-	- -
12	-	-		-				-	-	-

Note: The adjustment is for ADL, cognition, age, gender and length of stay.

The differences in facility QOL scores can be due to differences in facility performance, but they can also be attributed to resident case mix and the differences in the way interviewers performed. Table 6.4 presents a comparison of the ability of various components of QOL to discriminate among facilities using raw, non-adjusted measures and adjusted measures. The numbers in the table represent relative variance and p-values associated with a corresponding random factor. Higher relative variance and lower p-values indicate better chances to discriminate the analyzed entities (i.e., facilities, interviewers). When comparing p-values, correction for multiple comparisons has to be applied and the threshold p-value of 0.005 instead of 0.05 should be used. Prior to risk-adjustment six domains out often show a p-value less than 0.005, which corresponds to a relative variance greater than 3%. In one case (Spiritual Well-being) this value is close to 6%. Two more domains (Meaningful Activity, Autonomy) fall in the range of p-values from 0.005 to 0.05. When risk-adjustment was applied for patients' characteristics (covariates included into the model), no dramatic changes were found. After adjustment, the relative variance of the Facility factor increased in five cases and decreased in five cases. Six domains demonstrated p-values less than 0.005 and three between 0.005 and 0.05. Two domains did not show any significant contribution of the Facility factor. As a result of risk-adjustment, the overall level of significance associated with the same relative variance dropped because of the degrees of freedom consumed by risk-adjustment. For example, the factor Facility prior to adjustment was able to explain 2.2% of total variance of autonomy with p=0.007. After adjustment the relative variance increased to 2.4% but the corresponding significance became slightly worse (p=0.009).

Table 6.4. Ability of Various Components of QOL to Differentiate Facilities
(Non-adjusted and Adjusted for Residents' Characteristics)

	Non-adjusted		Adjusted for residents' characteristics	
	% total variance explained by Facility	P value	% total variance explained by Facility	P value
Comfort	3.1	0.0001	5.1	0.0001
Functional Competence	1.1	0.103	0.7	0.245
Privacy	3.2	0.001	3.8	0.0001
Dignity	1.0	0.053	0.9	0.075
Meaningful Activity	2.5	0.008	2.7	0.008
Enjoyment	3.5	0.001	3.9	0.001
Relationships	4.1	0.0001	3.6	0.001
Security	3.0	0.002	1.7	0.046
Spiritual Well Being	5.9	0.0001	5.2	0.0001
Autonomy	2.2	0.007	2.4	0.009

The study design did not specify how interviewers were distributed across facilities. Consequently, some interviewers visited as few as two facilities and some visited as many as seven facilities. All interviewers were trained and demonstrated acceptable inter-rater reliability. Nevertheless, that fact does not exclude the possibility that observed differences between facilities might be partially related to differences among interviewers. To test this assumption an additional correction for the random Interviewer factor was applied (Table 6.5). Eight domains demonstrated a loss of relative variance; only two domains did not show any change. In four cases variance associated with Interviewers was even greater than variance associated with Facilities (comfort, privacy, meaningful activity, security). This trend was combined with the fewer degrees of freedom; therefore, just one Facility p value (domain, dignity) fell in the range 0.005 to 0.05 and no one domain had a significant p-value ($P<0.005$).

Table 6.5. Ability of Various Components of QOL to Differentiate Facilities
(Adjusted for Residents' Characteristics and Allocation of Interviewers)

	% total variance explained by Facility	P value	% total variance explained by Interviewer	P value
Comfort	2.9	0.034	5.1	0.001
Functional Competency	0.7	0.301	0.0	0.725
Privacy	2.1	0.232	2.8	0.195
Dignity	0.9	0.049	0.5	0.220
Meaningful Activities	1.6	0.173	2.7	0.059
Enjoyment	2.8	0.145	1.7	0.161
Relationships	2.8	0.155	1.1	0.685
Security	0.3	0.724	2.0	0.422
Spiritual Well Being	3.5	0.070	1.5	0.496
Autonomy	2.4	0.073	0.0	0.600

The effect of losing degrees of freedom can be illustrated by comparing models for autonomy (Tables 6.4 and 6.5). Table 6.4 shows that the model adjusted for patient characteristics explains 2.4% of total variance and the Facility factor has p-value equal to 0.009. Table 6.5 shows that the model with added Interviewer factor has the same relative variance of 2.4% associated with the Facility factor and zero variance explained by the Interviewer factor. Nevertheless significance of the Facility factor became much worse (0.073 instead of 0.009).

To explore whether QOL changes as residents stay longer, we compared the various domains scores by length of stay. As shown in Table 6.6, there is no substantial change in the domain scores over time.

Table 6.6. Effect of Length of NH Stay on Domain Scores at Wave 1

	Length of Stay (months)							
Domain	< 1 month	1-2	2- 3	3- 6	6-12	12-18	18-24	> 24 months
Comfort	3.04	3	2.9	3.06	2.96	3.03	3.14	3.03
Functional Competence	3.1	3.29	3.24	3.34	3.25	3.25	3.28	3.22
Privacy	3.35	3.38	3.37	3.36	3.38	3.28	3.27	3.35
Dignity	3.73	3.79	3.68	3.71	3.68	3.67	3.7	3.63
Meaningful Activity	2.74	2.65	2.45	2.72	2.73	2.62	2.75	2.72
Enjoyment	3.39	3.09	3.09	3.34	3.11	3.27	3.22	3.24
Individuality	2.87	2.6	2.67	2.81	2.87	2.83	2.7	2.87
Spiritual Well-being	3.12	3.07	3.01	3.08	3.09	3.13	3.15	3.23
Security	3.61	3.58	3.47	3.43	3.39	3.4	3.4	3.33
Autonomy	3.53	3.38	3.2	3.38	3.34	3.23	3.27	3.28
Relationships	3.19	2.95	2.9	3.02	3.09	3.07	3.01	3.12
# of res.	40	108	93	200	290	214	174	867

Wave 2 Findings

Table 6.7 shows the characteristics of facilities in Wave 2. Because a different sampling scheme was employed, the characteristics differ from those of Wave 1. The facilities are generally larger and have many fewer single rooms. The staffing levels are generally comparable. Fewer facilities are proprietary and the proportion of Medicaid patients is slightly less. Many more residents are cognitively impaired.

Table 6.7 Wave 2 Facility Characteristics

Characteristics	**Sample facility rates**	**N**	**Data Source**
	Mean (range)		
No. of licensed beds	156.79 (67-559)	56	mailed survey
% of private rooms	14.24% (0-100%)	57	mailed survey
Staff/resident ratio	0.66 (0.41-.21)	61	OSCAR
% Proprietary	57.40%	57	mailed survey
% of Medicaid	63% (0-98%)	55	mailed survey
% of more cognitively impaired residents (MDS score of 3-5)	68% (38-91%)	60	MDS

The same pattern of variation in performance across QOL domains was seen as that in Wave 1 As Table 6.8 shows. there were fewer extreme cases; but, in general, facilities tended to be primarily positive or negative.

Table 6.8. Facility Risk-adjusted Performance Summary at Wave 2

NF	Domain										
	Comfort	Functional Competence	Privacy	Dignity	Meaningful Activities	Enjoyment	Individ-uality	Relation-ships	Security	Spiritual Well-being	Autonomy
1-01	+	-									+
1-02			+								
1-03	+		+	+		++	+	+			+
1-04		+	+	+		+	++	+	+	+	
1-05					+			++			
1-06		-					-				
1-07	-									-	--
1-08				+	+					+	-
1-09			-	-					-	+	-
1-10	-	+		+							
1-11				--	-		--	-	--		-
1-12		+							-	+	
2-13	-			--		-	-	-	-		
2-14	-	-		--	+		-		--		
2-15									-		
2-16		-			++			+			
2-17				+		+	+	+	++	+	
2-18							+			-	+
2-19	-			-						+	
2-20			-		+		-				
2-21	+	+		+				-	+	-	++
2-22	+			+	+		++	++			+
2-23		+								-	
2-24	-			-				-	-	-	
2-25			+				-	-		-	
3-31	-	-		-	-	--			-		
3-32	+	-			--	-					
3-33											
3-34	+							-		-	
3-35					+				+		+
3-36	+									+	+
3-37											
3-38	++	++	+	+	+		+		+		
3-39				+					-		

.

Table 6.8, cont'd

	Comfort	Functional Competence	Privacy	Dignity	Meaningful Activities	Enjoyment	Individuality	Relationships	Security	Spiritual Well-being	Autonomy
3-40		-				--	-	-		-	
3-41									+	-	
3-42	++	+	+		-	--		-		--	
4-70			+			-		+			
4-71	-					-			-		
4-72	-				-					-	+
4-73		--	--			-	-	-			--
4-74	+				+	+		++		+	
4-75								++			
4-76				-		+				+	
4-77				+			+				
4-78					-						
4-79	+		+			+			-	+	
4-80		+	+								+
4-81						+					
6-50	+				+	++	+	+	+	-	
6-51					+						
6-52											-
6-53								-	+	+	
6-54					-	-	-	-			
6-55											-
6-56			-			-		-			
6-57	-	+	-		-						
6-58	-		-		-			+			--
6-59						+					
6-60			---		-						-
6-61	+	-								+	--

The sample in Wave 2 included more short-stay residents and permitted closer inspection of the effects of lengths of stay. The same pattern of domain scores over time seen in Wave 1 is found in Wave 2. As shown in Table 6.9, residents' domain scores did not change appreciably with their length of stay.

Table 6.9 Effect of Length of NH Stay on Domain Scores—Wave 2

	Length of Stay (months)							
Domain	< 1 month	1 ~ 2	2 ~ 3	3 ~ 6	6 ~ 12	12 ~ 18	18 ~ 24	> 24 months
Comfort	2.93	2.93	2.85	2.99	2.99	3.06	2.95	2.98
Functional Competence	3.07	2.95	2.89	3.19	3.1	3.18	3.16	3.04
Privacy	3.16	3.26	3.15	3.17	3.16	3.24	3.18	3.09
Dignity	3.57	3.56	3.5	3.59	3.65	3.64	3.58	3.58
Meaningful Activity	2.36	2.49	2.59	2.78	2.71	2.83	2.82	2.76
Enjoyment	2.98	3.02	3.03	3.18	3.04	3.17	3.02	3.14
Individuality	2.9	3.05	3.06	3.07	3.09	3.21	3.08	3.12
Spiritual Well-being	2.84	2.75	3.06	2.99	3.08	3.11	3.05	3.12
Security	3.45	3.4	3.4	3.38	3.32	3.36	3.29	3.32
Autonomy	3.17	3.23	3.12	3.21	3.2	3.25	3.22	3.22
Relationships	2.82	2.89	2.91	2.9	3.01	2.99	3.01	3.02
# of residents (1662)	145	104	86	202	255	185	153	532

Facility Effect Size

An important question is the sample size needed to produce significantly different facility scores. To estimate the effect of different sample sizes, we used a method described by Dupont and Plummer (Dupont & Plummer, 1990). This method, which produces results that are in close agreement with those of Pearson and Hartley, uses the relationship of the between-group difference and the within-group difference (Pearson & Hartley, 1970). Because these data could be used to describe facility performance, special care must be taken to assure that false positives (Type II errors) are avoided. Hence a conservative alpha value should be used. Table 6.10 uses Wave 1 data to show the estimated sample sizes needed to detect differences between two facilities that exceed one within-group standard deviation, but the choice of a one standard deviation threshold difference is arbitrary. In the first case, the risk of declaring a difference due simply to chance variation in sampling (alpha) is less than 5% and the chance of declaring no difference when there really might be one (beta or 1-power) is 20%. In the second case the alpha

level is set at 1% and the beta is set at 10%. The sample size varies among the domains. For the first case, it runs from 7 (spiritual well-being) to 12 (dignity and functional competence). For the second, more stringent case, it runs from 13 to 23. Thus, a sample size of 23 per nursing home would permit a reasonable comparison across facilities. This number represents the number of nursing home residents who responded to an adequate number of the questions about QOL. Presumably a somewhat larger sample would have to be approached to net this total. If smaller units within a facility were to be compared, then the necessary sample size would increase considerably.

When the same calculations are made using Wave 2 data (Table 6.11), the sample sizes needed are comparable. For the less stringent case, it ranges from 6 to 13, and for the more stringent case from 13 to 25.

Table 6.10. Sample Size Needed to Detect A Mean Difference Across Facilities: Equal to Different Levels of Between-Groups Standard Deviation at Wave 1

Domain	Between Groups	Within Groups	Observed between-group SD expressed in units of within-group SD	Sample Size	
				Alpha = .05 Power = .80	Alpha = .01 Power = .90
Comfort	0.88	0.61	1.45	9	16
Functional Competence	0.86	0.73	1.18	12	23
Privacy	0.87	0.62	1.40	9	17
Dignity	0.50	0.43	1.18	12	23
Meaningful Activity	0.92	0.71	1.29	10	19
Enjoyment	1.00	0.72	1.39	9	17
Individuality	1.02	0.68	1.50	8	15
Relationships	0.73	0.55	1.33	10	18
Security	1.13	0.68	1.65	7	13
Spiritual Well being	0.82	0.64	1.28	11	20
Autonomy	0.88	0.61	1.45	9	16

Table 6.11. Sample Size Needed to Detect A Mean Difference Across Facilities: Equal to Different Levels of Between-Groups Standard Deviation at Wave 2

Domain	Between Groups	Within Groups	Observed between-group SD expressed in units of within-group SD	Sample Size	
				Alpha = .05 Power = .80	Alpha = .01 Power = .90
Comfort	0.84	0.61	1.37	9	17
Functional Competence	1.07	0.83	1.29	10	19
Privacy	1.23	0.68	1.81	6	11
Dignity	0.58	0.52	1.13	13	25
Meaningful Activity	0.96	0.75	1.28	11	20
Enjoyment	1.15	0.78	1.48	8	15
Individuality	0.87	0.68	1.28	11	20
Relationships	0.93	0.70	1.33	10	18
Security	0.78	0.57	1.38	9	17
Spiritual Well-being	1.15	0.73	1.58	7	13
Autonomy	0.83	0.66	1.27	11	20

Discussion

These analyses suggest that it is possible to create facility-level scores for QOL and to compare facilities on this basis. At this point, we have used 10 separate QOL scores for each domain. A further step would be to create a weighted score to establish a single QOL score for each facility. This would render comparisons simpler, but might, in fact, obscure differences of interest, especially when the nursing home profile has both positive and negative deviations, a situation that occurred 30% of the time in our data. Moreover, from the perspective of the quality assurance requirements, nursing homes are accountable for outcomes in specific areas such as dignity, rather than the general construct QOL.

Despite the unusual configuration of the sample, the facility scores did appear to identify facilities with better and worse QOL. In general, the performance pattern showed consistency. That is, facilities that were below the mean in one area tended to show a similar pattern for other domains, which may result from correlated facility-level domain scores. The ANOVA test in

Table 6.4 revealed significant differences between facilities in eight domains out of ten, indicating that among all studied facilities there is at least one pair with significantly different QOL scores on eight domains. It makes them the preferable choice in comparing facilities in practical applications when applying the six-sigma methodology (i.e., more than two standard deviations from the mean value) to find outliers. Two domains (functional competence and dignity) failed to show any significance, indicating that all observed differences between facilities on these domains were likely to appear by chance. Although the practical value of these domains is very limited when comparing facilities, they might be useful in longitudinal studies organized to test the effects of facility-level organizational interventions. One concern with using domain averages is that if these average scores are near the upper or lower ends of the possible range, a ceiling or floor effect could result, and we would not be able to detect any positive or negative deviations. However, the results in Table 6.2 show that this potential threat did not occur. None of the average facility-level domain scores approached either end of the possible range.

The general lack of statistically significant relationships between QOL scores and nursing home characteristics may be attributed in part to the limited sample size. The sample itself is atypical of the general distribution of nursing homes in the United States. By design, a disproportionately large number of rural homes were included. The lower proportion of proprietary homes may reflect the effect of this sampling, as well as refusals to participate.

Some limitations of this study should be acknowledged. We have used z-scoring techniques to simplify the data presentation, but these transformations could create apparently large differences where the actual differences in scores are small. Care must be taken in interpreting the relative performance of any facility. However, if attention is focused on patterns

across domains, the risk of over-interpretation is minimized. The sample used for this study was not intended to be nationally representative. The emphasis here was on developing a method for aggregating resident-level QOL scores to the facility level. Further work is needed to see how well this approach can discriminate among a larger, more representative sample of facilities.

While analyzing the possibility of various QOL domains to discriminate facilities, we found that difference in patient characteristics between facilities (the most commonly used reason for adjustment) changed the resolution of the comparison very little. A much greater source of variation was attributable to the measurement process (i.e., allocation of interviewers to facilities). Because the study had not been planned to examine this effect, we had to rely on a post hoc analysis of this observational study with a very unbalanced design and a sample size that was insufficient to test this effect. Nevertheless, the importance of the measurement process was demonstrated by comparing the relative variances associated with two random factors. (Interviewer and Facility factors were found to have comparable variance.) If confirmed, this finding would affect the process of comparing QOL between facilities. Such studies should be designed to prevent confounding the Facility and Interviewer factors and to allow separating their effects.

This study shows that it is possible to distinguish resident reports of QOL among nursing facilities. Even though the amount of overall variance explain by facilities is small compared to individual-level factors, it is important. Whereas resident characteristics are critical in predicting QOL, some portion of the variance remains under the control of facilities. They thus can be held accountable for this important aspect of nursing home life, if appropriate case-mix adjustments are made in the analysis of the results. The specific pattern of strengths and weakness across the various domains may be especially informative for quality improvement efforts.

References

Braun, B. I. (1991). The effect of nursing home quality on patient outcome. *Journal of the American Geriatrics Society*, 39(3), 329-388.

Davis, M. A. (1991). On nursing home quality: A review and analysis. *Medical Care Review*, 48(2), 129-166.

Dupont, W. D., & Plummer, W. D. (1990). Power and sample size calculations: A Review and computer program. Controlled Clinical Trials., 11, 116-128.

Mukamel, D. B. (1997). Risk-Adjusted Outcome Measures and Quality of Care in Nursing Homes. Medical Care, 35(4), 367-385.

Pearson, E. S., & Hartley, H. O. (1970). Biometrika Tables for Statisticians, 3rd Edition, Volume 1. Cambridge, UK:: Cambridge University Press.

Phillips, C., D., Hawes, C., Mor, V., Fries, B. E., Morris, J. N., & Nennsstiel, M. (1996). Facility and area variation affecting the use of physical restraints in nursing homes. Medical Care, 34(11), 1149-1162.

Porell, F., & Caro, F. G. (1998). Facility-level outcome performance measures for nursing homes. *The Gerontologist*, 38(6), 665-683.